MAPPING OF THE ALTITUDINAL AND TOPOGRAPHICAL PATTERNS OF MOUNTAINOUS VEGETATION IN NORTHEAST JAPAN

Ram C. Sharma (1), Hidetake Hirayama (2), Mizuki Tomita (1) and Keitarou Hara (1)

[1]Department of Informatics, Tokyo University of Information Sciences, 4-1 Onaridai, Wakaba-ku, Chiba 265-8501, Japan

[2] Graduate School of Tokyo University of Information Sciences, 4-1 Onaridai, Wakaba-ku, Chiba 265-8501, Japan

Email: sharma@rsch.tuis.ac.jp; h17002hh@edu.tuis.ac.jp; tomita@rsch.tuis.ac.jp; hara@rsch.tuis.ac.jp

KEY WORDS: Vegetation, Topography, Landsat 8, Hakkoda Mountains, Random Forests

ABSTRACT: The mountainous habitats are one of the climate sensitive zones. In the mountainous regions, distinct altitudinal and topographical patterns of vegetation are formed by the interaction of varying environmental factors such as soil composition, temperature, precipitation, humidity, and solar radiation. Satellite remote sensing is a suitable technology for wide area monitoring of vegetation patterns to better understand the climatic response of vegetation. The major objectives of this research were to assess the potential of machine learning technique to explain the altitudinal and topographic response of mountainous vegetation patterns, and to map major vegetation types in the Hakkoda Mountains. The Hakkoda Mountains represent a typical mountainous ecosystem in the cool temperate climatic region of north eastern Japan. The dominant plants of the region include *Fagus crenata*, *Quercus crispula*, *Betula ermanii*, *Abies mariesii*, and *Pinus pumila*. All Landsat 8 data of 2018 available for the study area were processed (cloud masking, mosaicking, multi-temporal compositing) and large number of spectral and spectral-indices features were generated. We also generated the topographical features (elevation, slope, aspect, and hill shade) from the Shuttle Radar Topography Mission (SRTM) product. Random Forests algorithm was employed for assessing feature importance and classification of vegetation types. The results were validated with the ground truth data prepared in the research. The altitudinal and topographical response of the vegetation patterns has been discussed, and vegetation map of the Hakkoda Mountains has been presented. The methodology described in the research is expected to be applicable in understanding mountainous vegetation patterns of other regions as well.

1. INTRODUCTION

The mountainous habitats are one of the climate sensitive zones. Satellite remote sensing is a suitable technology for wide area monitoring of vegetation patterns and better understanding the climatic response of vegetation.

Landsat 8 is one of the satellites appropriate for this purpose. Landsat 8 has been collecting global scale, 15-100m spatial resolution, multi-spectral images with a standard 16-day repeat cycle since 2013 (Irons et al., 2012; Roy et al., 2014). Landsat 8 consists of two sensors, Operational Land Imager (OLI) and Thermal InfraRed Sensor (TIRS); and collects image data for nine visible-shortwave bands and two thermal bands (www1). The geo-registration of Tier 1 scenes is consistent within prescribed image-to-image tolerances of $\leqq$ 12-meter radial root mean square error (www1). The list of Landsat 8 (OLI and TIRS) bands relevant to vegetation mapping has been shown in Table 1.

Table 1. List of Landsat 8 (OLI and TIRS) bands relevant to land cover and vegetation mapping.

Bands	Wavelength (micrometers)	Resolution (meters)
Band 2 - Blue	0.45-0.51	30
Band 3 - Green	0.53-0.59	30
Band 4 - Red	0.64-0.67	30
Band 5 - Near Infrared (NIR)	0.85-0.88	30
Band 6 - Mid Infrared (MIR)	1.57-1.65	30
Band 7 – Shortwave Infrared (SWIR)	2.11-2.29	30
Band 8 - Panchromatic	0.50-0.68	15
Band 11 - Thermal Infrared (TIRS) 2	11.50-12.51	100

The spectral indices, arithmetic combination of the spectral bands, are also crucial to the vegetation mapping. A list of multispectral indices relevant to vegetation mapping has been shown in Table 2.

Table 2. List of multispectral indices relevant to vegetation mapping.

| Multispectral indices	References

Green Red Vegetation Index (GRVI)	Falkowski et al., 2005
Normalized Difference Vegetation Index (NDVI)	Rouse et al., 1974
Land Surface Water Index (LSWI)	Xiao et al., 2004
Normalized Difference Built-up Index (NDBI)	Zha et al., 2003
Normalized Difference Snow Index (NDSI)	Hall et al., 2002
Normalized Difference Water Index (NDWI)	McFeeters, 1996

In mountainous regions, distinct altitudinal and topographical patterns of vegetation are formed by the interaction of varying environmental factors such as soil composition, temperature, precipitation, humidity, and solar radiation. The zonation of vegetation with respect to altitudinal and topographical gradients has long been recognized (Ohsawa, 1984; Sakai and Ohsawa, 1994). Therefore, it is interesting to know to what extent the altitudinal and topographic variation of the vegetation patterns can be explained by machine learning techniques. The major objectives of this research were to quantitatively assess the potential of machine learning technique to explain the altitudinal and topographic response of mountainous vegetation patterns, and to map major vegetation types using satellite data.

2.   MATERIALS AND METHODS
2.1 Study Area

This research was conducted in the Hakkoda Mountains, which represent a typical mountainous ecosystem in the cool temperate climatic region of north eastern Japan. The dominant plants of the region include *Fagus crenata*, *Quercus crispula*, *Betula ermanii*, *Abies mariesii*, and *Pinus pumila*. The location of the study area is shown in Figure 1.
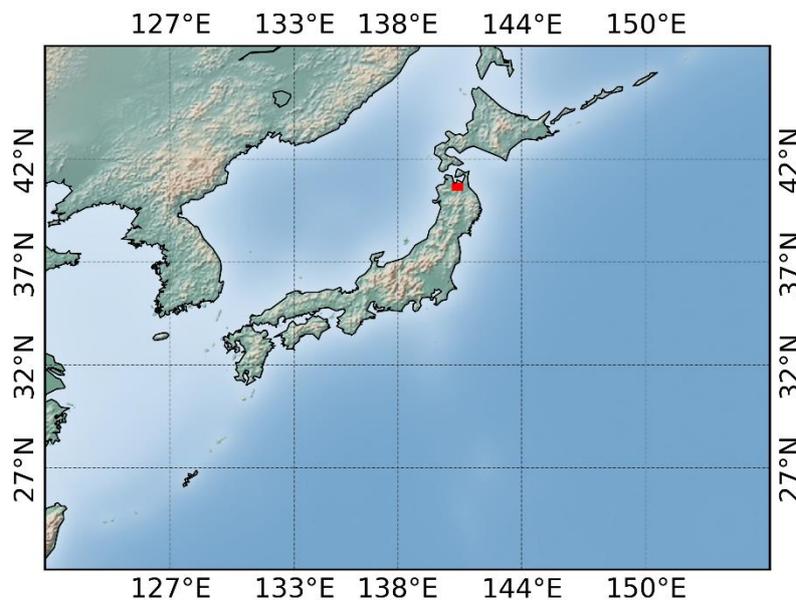
Figure 1. Location of the study area, the Hakkoda Mountains, shown by red polygon

The vegetation patterns in one of the Hakkoda Mountains is shown in Figure 2.

Figure 2. One of the Hakkoda Mountains showing different patterns of vegetation

2.2 Preparation of ground truth data

We collected ground truth data with reference to vegetation survey map (www2), Google Earth (www3), and field survey. The distribution of ground truth data in the study area is shown in Figure 3. For each vegetation class, 250 sample points were prepared as the ground truth data.
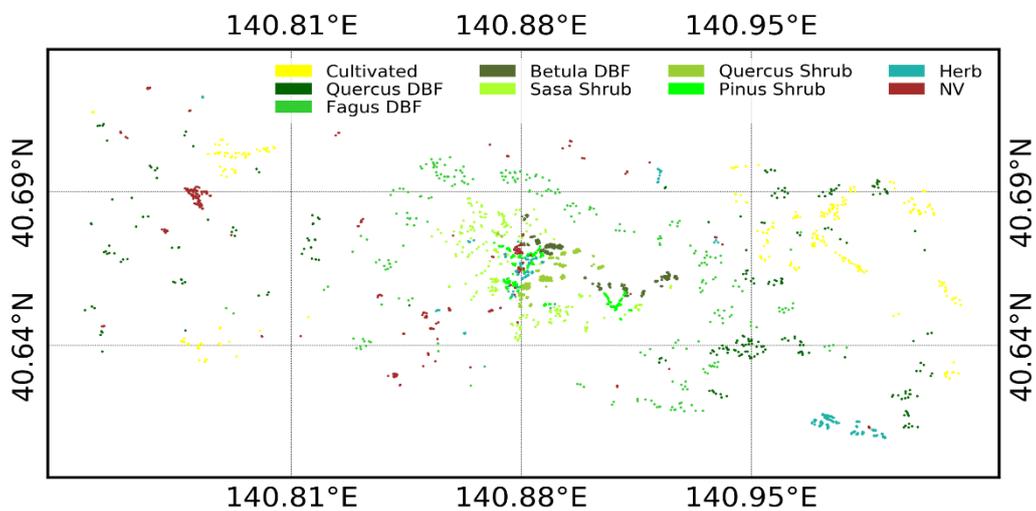
Figure 3. Distribution of ground truth data (prepared in the research) over the study area

2.3 Processing of satellite data

We used all Landsat-8 scenes (in total 56) under the Tier 1 collections available for the study area in 2018. Landsat-8 data were converted into top-of-atmosphere (TOA) spectral reflectance using the rescaling coefficients. The cloud cover of the scenes varied from 4.85 – 93.47%. The clouds were removed by using separate quality assessment (QA) band information. The spectral bands (Table 1) were extracted and the spectral indices (Table 2) were calculated. The spectral/spectral-indices images were composited using percentile based method (Sharma et al., 2017). Table 4 shows the list of Landsat 8 features generated in the research.

3

Table 4. Description of Landsat 8 features (spectral and spectral-indices) at different percentiles (p) generated for the research.

Landsat 8 features							
blue p0	green p90	nir p70	swir p50	tir p30	grvi p10	ndwi p100	ndbi p80
blue p10	green p100	nir p80	swir p60	tir p40	grvi p20	lswi p0	ndbi p90
blue p20	red p0	nir p90	swir p70	tir p50	grvi p30	lswi p10	ndbi p100
blue p30	red p10	nir p100	swir p80	tir p60	grvi p40	lswi p20	ndsi p0
blue p40	red p20	mir p0	swir p90	tir p70	grvi p50	lswi p30	ndsi p10
blue p50	red p30	mir p10	swir p100	tir p80	grvi p60	lswi p40	ndsi p20
blue p60	red p40	mir p20	pan p0	tir p90	grvi p70	lswi p50	ndsi p30
blue p70	red p50	mir p30	pan p10	tir p100	grvi p80	lswi p60	ndsi p40
blue p80	red p60	mir p40	pan p20	ndvi p0	grvi p90	lswi p70	ndsi p50
blue p90	red p70	mir p50	pan p30	ndvi p10	grvi p100	lswi p80	ndsi p60
blue p100	red p80	mir p60	pan p40	ndvi p20	ndwi p0	lswi p90	ndsi p70
green p0	red p90	mir p70	pan p50	ndvi p30	ndwi p10	lswi p100	ndsi p80
green p10	red p100	mir p80	pan p60	ndvi p40	ndwi p20	ndbi p0	ndsi p90
green p20	nir p0	mir p90	pan p70	ndvi p50	ndwi p30	ndbi p10	ndsi p100
green p30	nir p10	mir p100	pan p80	ndvi p60	ndwi p40	ndbi p20	
green p40	nir p20	swir p0	pan p90	ndvi p70	ndwi p50	ndbi p30	
green p50	nir p30	swir p10	pan p100	ndvi p80	ndwi p60	ndbi p40	
green p60	nir p40	swir p20	tir p0	ndvi p90	ndwi p70	ndbi p50	
green p70	nir p50	swir p30	tir p10	ndvi p100	ndwi p80	ndbi p60	
green p80	nir p60	swir p40	tir p20	grvi p0	ndwi p90	ndbi p70	

We used Shuttle Radar Topography Mission (SRTM) V3 product at a resolution of 30m (Farr et al., 2007) for generating the topographic features (elevation, slope, aspect, and hill shade).

Altogether, 158 features (154 from Landsat 8 and 4 from SRTM) were generated.

2.4   Machine learning

We employed Random Forests classifier for the supervised classification of the Landsat 8 and Topographic features with the support of ground truth data prepared in the research. Random forests are the random decision forests that operates by constructing a multitude of decision trees to overcome overfitting problem of the decision trees (Ho, 1995; Breiman, 2001). The feature importance were identified with inbuilt feature importance module available in the Random Forests algorithm (Breiman, 2001). Out of 250 ground truth points, 150 points were used as the training data, whereas remaining 100 points were used as the validation data. The results were evaluated by feature importance, confusion matrix, overall accuracy, and kappa coefficient obtained with the validation data.
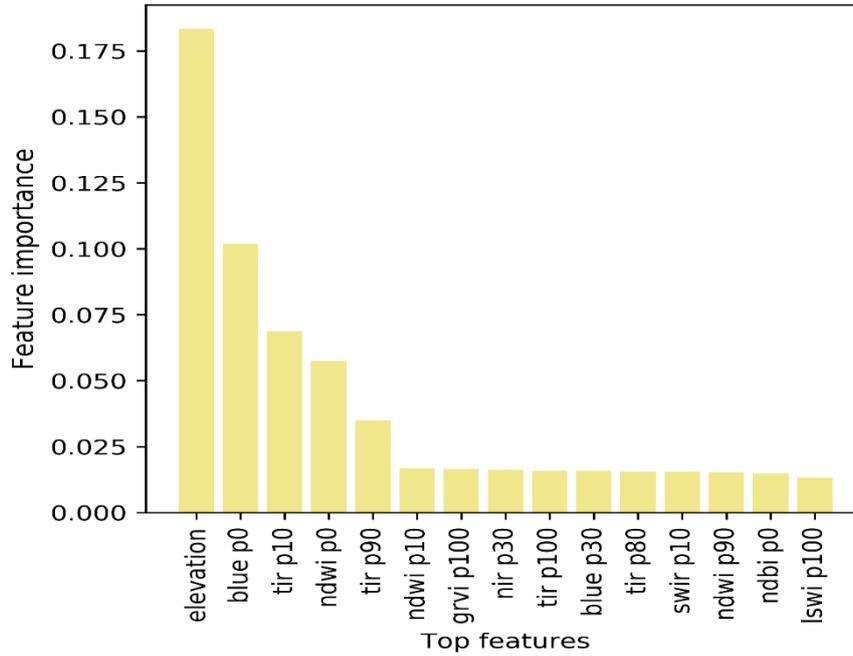
3. RESULTS

The importance of different set of features (Landsat 8, Topographic, and Landsat 8 + Topographic) obtained from the Random Forests classifier are shown in Figures 4-6. Overall, the topographic (elevation) feature was found to be most determining factor for explaining the vegetation patterns in the study area (Figure 6). Among the topographic features, aspect was found to be second most explanatory factor after the elevation (Figure 5). Both the spectral and spectral-indices were found to be important for vegetation classification (Figure 4).
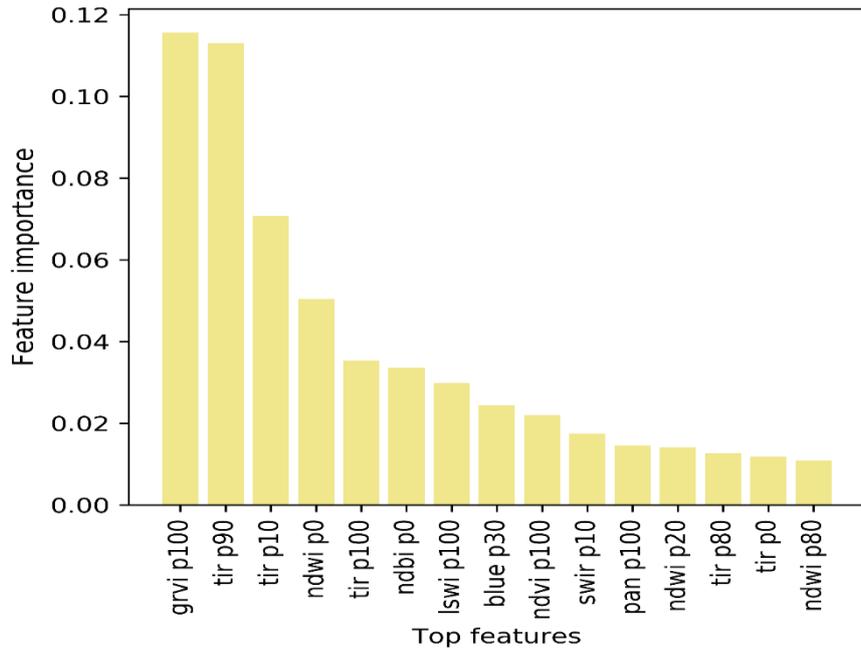
**Figure 4**. Importance of Landsat 8 features
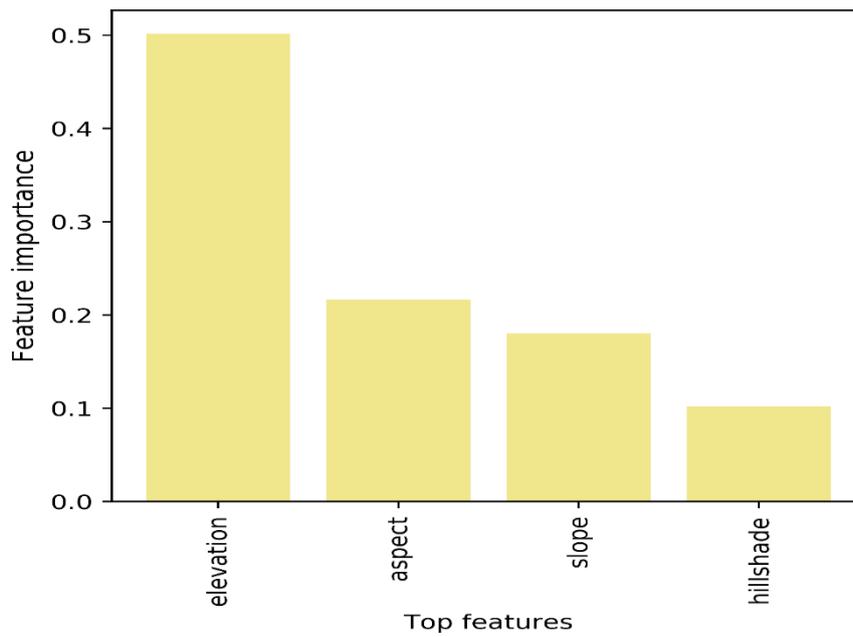
**Figure 5**. Importance of Topographic features

**Figure 6**. Importance of Landsat 8 + Topographic features

The validation results also showed importance (Overall accuracy = 0.72; Kappa = 0.69) of the topographic features for classification of vegetation types in the study area (Figure 7). Moreover, the combination of topographic features with the Landsat 8 features (Figure 9) provided better accuracy (Overall accuracy = 0.91; Kappa = 0.90) than using Landsat 8 features (Figure 8) alone (Overall accuracy = 0.89; Kappa = 0.88).
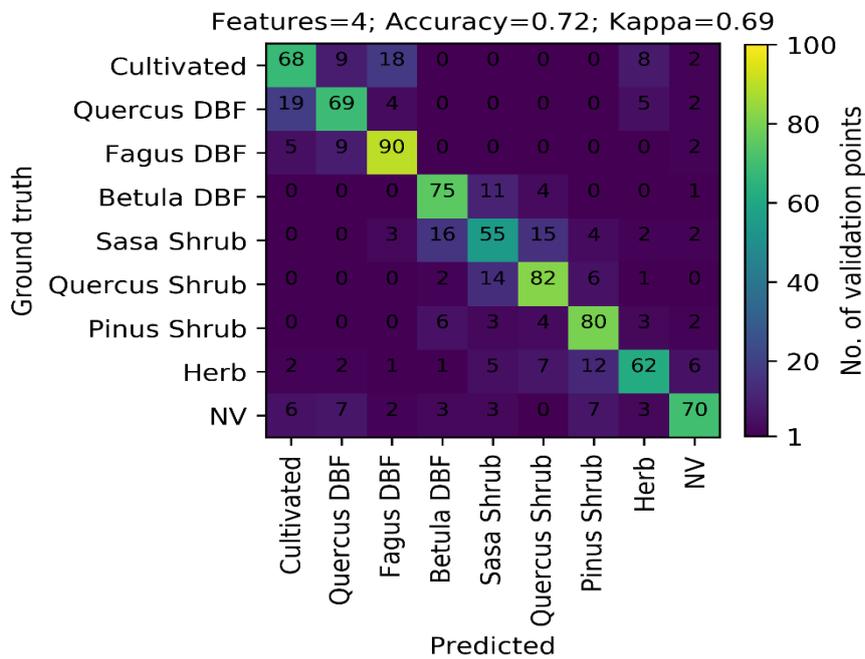
**Figure 7**. Validation results using Topographic features
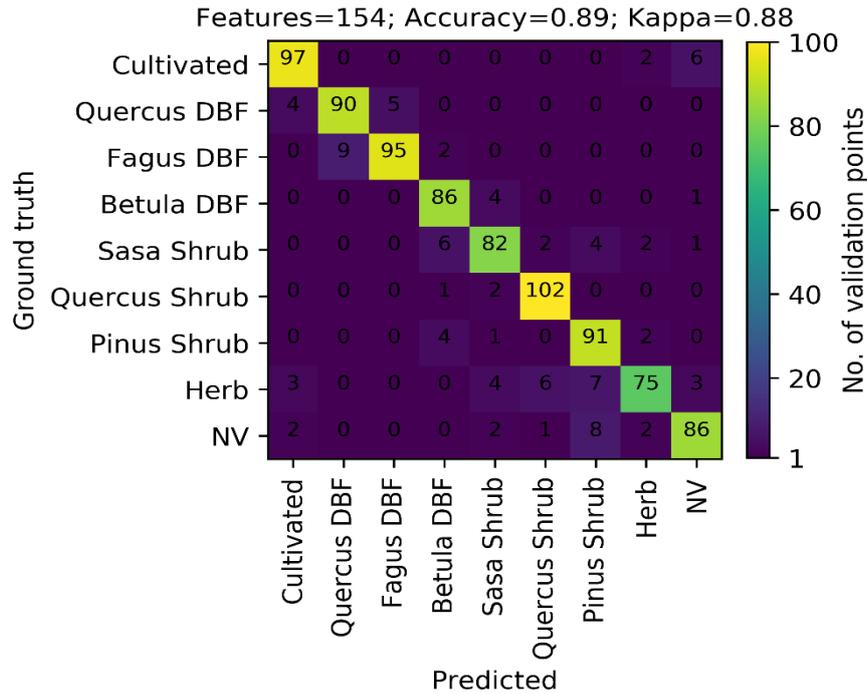
Figure 8. Validation results using Landsat 8 features
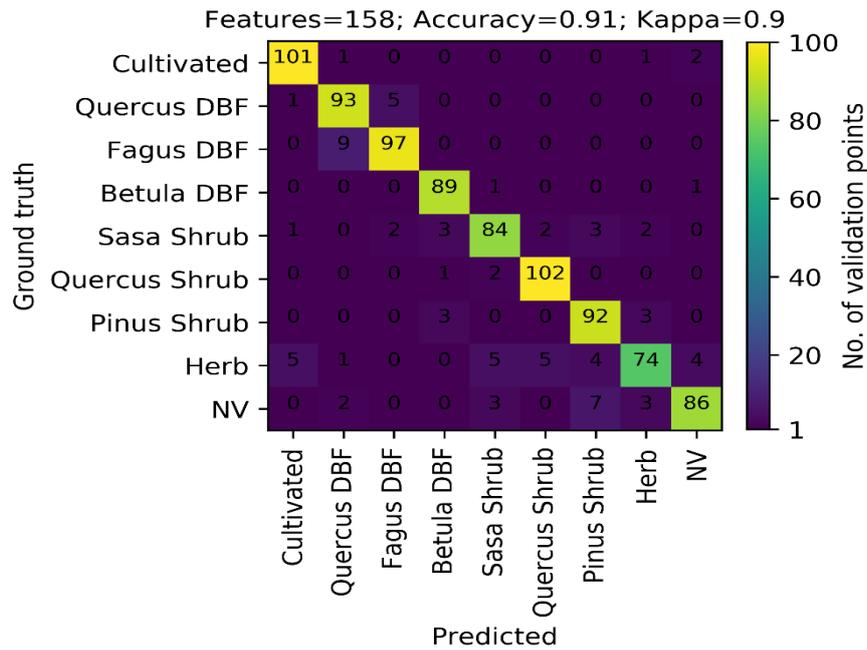
Figure 9. Validation results using Landsat 8 + Topographic features

The vegetation map of the Hakkoda Mountains produced by using Landsat 8 and SRTM data is shown in Figure 10. The large number of multi-temporal (spectral and spectral-indices) and topographic (elevation, aspect, slope and hill shade) features has resulted vegetation map of high accuracy in the research.

7
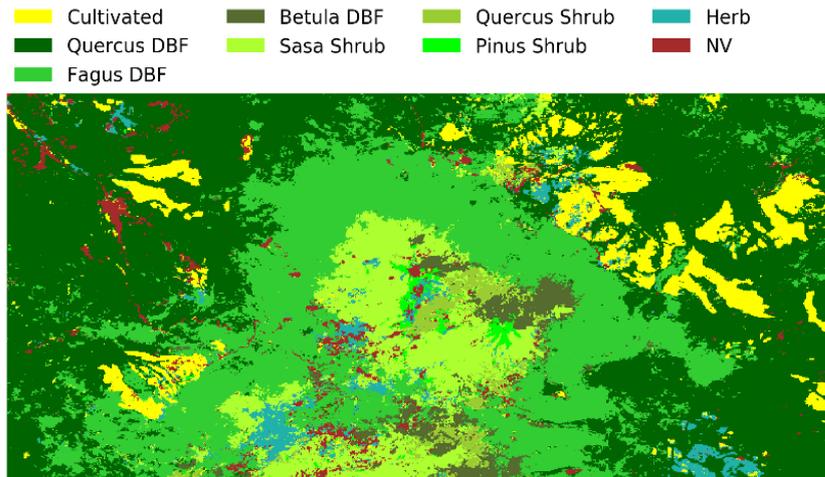
Figure 10. Vegetation map of the study area

4. CONCLUSION

In this research, the machine learning technique (Random Forests classifier) showed the potential of explaining the altitudinal and topographic response of vegetation patterns in Hakkoda Mountains. The topographic (elevation) feature was found to be most determining factor for explaining the vegetation patterns in the study area. We produced vegetation map with high accuracy by combining topographic features with Landsat 8 based spectral and spectral-indices features. The methodology described in the research is expected to be applicable in understanding mountainous vegetation patterns of other regions as well.

Acknowledgements

References

Breiman, L., 2001. Random forests. *Machine learning*, *45*(1), pp.5-32.

Falkowski, M.J., Gessler, P.E., Morgan, P., Hudak, A.T. and Smith, A.M., 2005. Characterizing and mapping forest fire fuels using ASTER imagery and gradient modeling. *Forest Ecology and Management*, *217*(2-3), pp.129-146.

Farr, T.G., Rosen, P.A., Caro, E., Crippen, R., Duren, R., Hensley, S., Kobrick, M., Paller, M., Rodriguez, E., Roth, L. and Seal, D., 2007. The shuttle radar topography mission. *Reviews of geophysics*, *45*(2).

Hall, D.K., Riggs, G.A., Salomonson, V.V., DiGirolamo, N.E. and Bayr, K.J., 2002. MODIS snow-cover products. *Remote sensing of Environment*, *83*(1-2), pp.181-194.

Ho, T.K., 1995, August. Random decision forests. In *Proceedings of 3rd international conference on document analysis and recognition* (Vol. 1, pp. 278-282). IEEE.

Irons, J.R., Dwyer, J.L., Barsi, J.A., 2012. The next Landsat satellite: The Landsat Data Continuity Mission. Remote Sensing of Environment 122, 11–21. https://doi.org/10.1016/j.rse.2011.08.026

McFeeters, S.K., 1996. The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. *International journal of remote sensing*, *17*(7), pp.1425-1432.

Ohsawa, M., 1984. Differentiation of vegetation zones and species strategies in the subalpine region of Mt. Fuji. *Vegetatio*, *57*(1), pp.15-52.

Rouse Jr, J., Haas, R., Schell, J., Deering, D., 1974. Monitoring vegetation systems in the Great Plains with ERTS

Roy, D.P., Wulder, M.A., Loveland, T.R., C.E., W., Allen, R.G., Anderson, M.C., Helder, D., Irons, J.R., et al., 2014. Landsat-8: Science and product vision for terrestrial global change research. Remote Sensing of Environment 145, 154–172. https://doi.org/10.1016/j.rse.2014.02.001

Sakai, A. and Ohsawa, M., 1994. Topographical pattern of the forest vegetation on a river basin in a warm-temperate hilly region, central Japan. *Ecological Research*, *9*(3), pp.269-280.

Sharma, R.C., Hara, K. and Hirayama, H., 2017. A machine learning and cross-validation approach for the discrimination of vegetation physiognomic types using satellite based multispectral and multitemporal data. *Scientifica*, *2017*.

Xiao, X., Hollinger, D., Aber, J., Goltz, M., Davidson, E.A., Zhang, Q. and Moore III, B., 2004. Satellite-based modeling of gross primary production in an evergreen needleleaf forest. *Remote sensing of environment*, *89*(4), pp.519-534.

Zha, Y., Gao, J. and Ni, S., 2003. Use of normalized difference built-up index in automatically mapping urban areas from TM imagery. *International journal of remote sensing*, *24*(3), pp.583-594.

www1. https://landsat.usgs.gov/landsat-8-data-users-handbook

www2. http://gis.biodic.go.jp/webgis/sc-025.html?kind=67

www3. http://earth.google.com/